MOTION BASED TARGET ACQUISITION AND EVALUATION IN AN ADAPTIVE MACHINE VISION SYSTEM

Presented at the Video Photogrammetry and Exploitation Conference, 11-12 May 1995, Washington, D.C.

Michael R. Blackburn, Ph.D.

D371, SPAWAR Systems Center, San Diego, CA 92152-7383
(619) 553-1904
mike@spawar.navy.mil

In this brief introduction to the topic there are sixteen points that we would like to make in reference to the problems of dynamic scene understanding and 3-D model extraction. We will first list these sixteen points, then take about a minute and a half to explain each one separately.

1. **Machine vision systems may achieve human-like perception most efficiently through progressive emulation of natural mechanisms of visual-motor control.**

2. **Motion is fundamental to all forms of natural perception.**

3. **Independent motion marks new targets, while induced motion provides information about the geometry of a static environment.**

4. **Target behavior is apparent primarily through an analysis of motion.**

5. **The geometries of natural vision systems facilitate processing of species relevant information.**

6. **Motion information can transform pattern information to achieve perceptual constancies.**

7. **Visual perception is an active process.**

8. **Reflex saccadic eye movements sample the environment.**

9. **Expectations drive search patterns over familiar targets.**

10. **Recognition is the verification of a prediction.**

11. **Acquisition and use of information are inseparable processes in natural intelligence.**

12. **Animals learn environmental correlations to satisfy internal needs.**

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 1995	**N/A**	**-**

4. TITLE AND SUBTITLE
Motion Based Target Acquisition and Evaluation in An Adaptive Machine Vision System

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Space and Naval Warfare Systems Center NCCOSC D746 Advanced Concepts Branch San Diego, CA 92152

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITORS ACRONYM(S)

11. SPONSOR/MONITORS REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
unclassified	**unclassified**	**unclassified**	**UU**	**24**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

13. **Machines can learn similarly if needs are appropriately defined and tested.**

14. **Machine learning, following biological precedent, requires a reflex base that responds to both internal and external events, sensor preprocessing for feature definitions, and association matrices between abstract representations of information from the sensor domains.**

15. **Neural networks, whether biological or artificial, self organize and select idiosyncratically relevant features for discrimination and prediction of environmental contingencies.**

16. **Recommendations and Summary of Machine vision at NRaD**

Now for some explanaton:

1) Machine vision systems should emulate natural mechanisms.

How to approach human-like perception without human liabilities?

What are the human liabilities?

> Unreliable - errors of omission, errors of commission,
> Unsuitable - slow, capacity limited
> Expensive - costs of training and maintenance,
> Fragile - costs of protection and repair.

What are the human assets?

> Adaptable - on-the-job learning,
> Available - many candidates for the job.

Why emulate biological mechanisms?

1. Natural mechanisms have proven successful and efficient.

2. A great deal is known of how they work.

3. Early fidelity to natural mechanisms may facilitate construction of higher order processes that depend upon them, and of which we yet are uncertain.

Advanced information processing systems such as man are phylogenetic consequences of simpler designs. Little is thrown away in the design of more advanced systems, rather, new capabilities are built up by the addition of neural controllers that interact with earlier

existing controllers. Figure 1 lists the relative complexity of the phylum, the evident nervous system advance at that stage, and the consequential new capabilities afforded.

UNCLASSIFIED

Phylum (family)	Example	Advance	Ability
Protozoa	Paramecium	• No nervous system	• Swim • Food discrimination
Coelenterata	Hydra Anemone Jellyfish	• Nerve nets • Statolith	• Spontaneity • Righting • Escape
Flatworms	Planaria	• Bilateral symmetry • Head ganglion • Commissures • Multisensors	• Kinesis • Taxis • Conditioning
Roundworms	Leech Earthworm	• Dorsal brain • Segmental ganglia • Central cord • Sense organs • Giant fibers	• Maze learning
Arthropods	Shrimp Lobster Spider Cockroach Bee	• Somatogastric and hear ganglia • Neurosecretory cells • Mixed sensor and motor nerves	• Startle response • Communication and social behavior • Migration • Selective predation
Mollusks	Snail Aplysia Octopus	• Brain lobes: – Centers for special senses – Motor and integration	• Short and long-term memory
Vertebrates	Fish Amphibia Reptile Bird	• Somatomotor reflexes in cord • Central chemo-receptors	• Sleep • Instrumental learning

Figure 1.

If we want to achieve the capabilities of man in an artificial system without the his/her limitations, we may do so by judicious emulation of the computational processes that subserve his intelligence.

2) Motion Analysis is fundamental

Motion dominates the processing of simpler organisms. In man there are specialized receptors for motion in cutaneous touch - the Meissner and Pacinian corpuscles, while other receptors - the Merkel and Ruffini - code static pressure (Vallbo, 1994); for the movement or change in stretch and tension of muscles there are the muscle spindle organs and the Golgi tendon organs, while and joint angle sensors code static position; in vision the rod photoreceptors and the magnocellular pathway are primarily involved in the processing of optic flow to visual motion, while the cone photoreceptors and the

parvocellular pathway are primarily concerned with the processing of pattern and color (van Essen and Maunsell, 1983). The functional difference between static and transient detectors is adaptation. Motion detectors rapidly adapt to conditions. (Muscle spindles adapt through active mechanisms involving the Gamma motor control of the intrafusal motor fibers.)

Motion is not something that was added to scene analysis, but it is what nature started with. Instead, pattern analysis was an addition to motion analysis.

Figure 2 is a mediolateral view of the human brain. All of the parts of the brain that can be identified in simpler species are located in progressively more central and more posterior regions.

UNCLASSIFIED

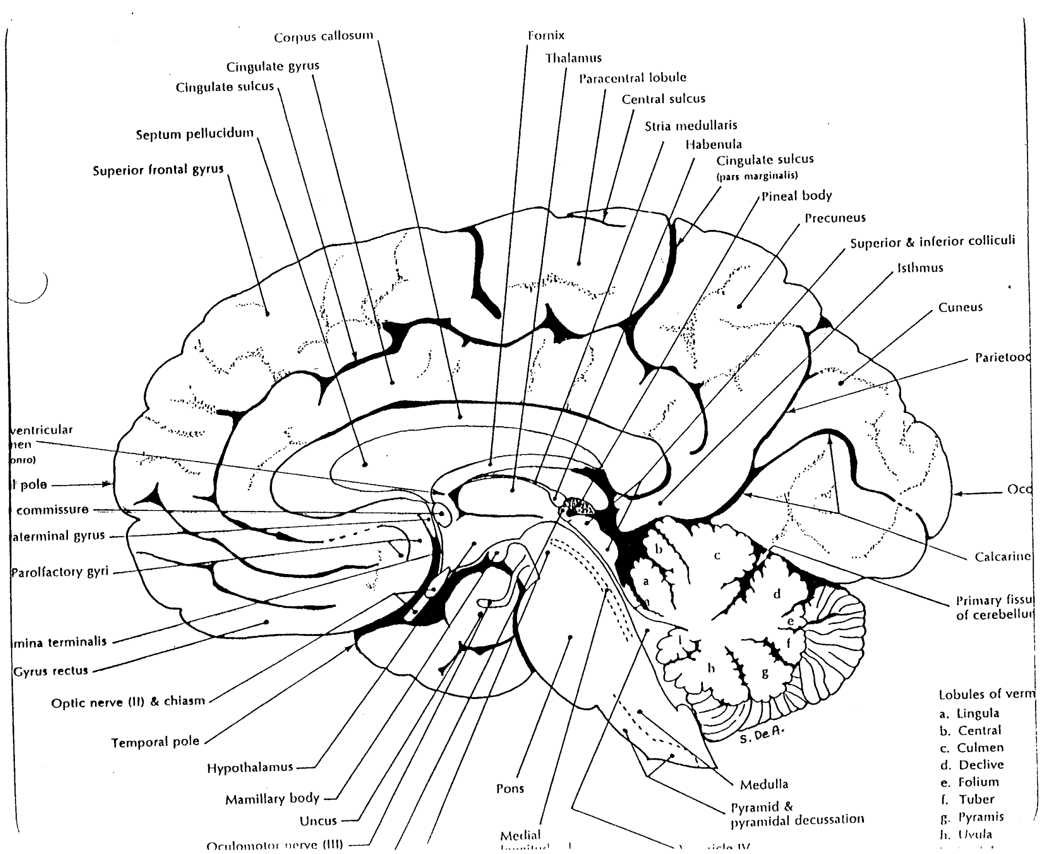


Figure 2.

3) Induced motion provides information about the geometry of a static environment.

Animals exploit their own ability to move by traversing the environment - creating local changes in pattern on their sensor fields.

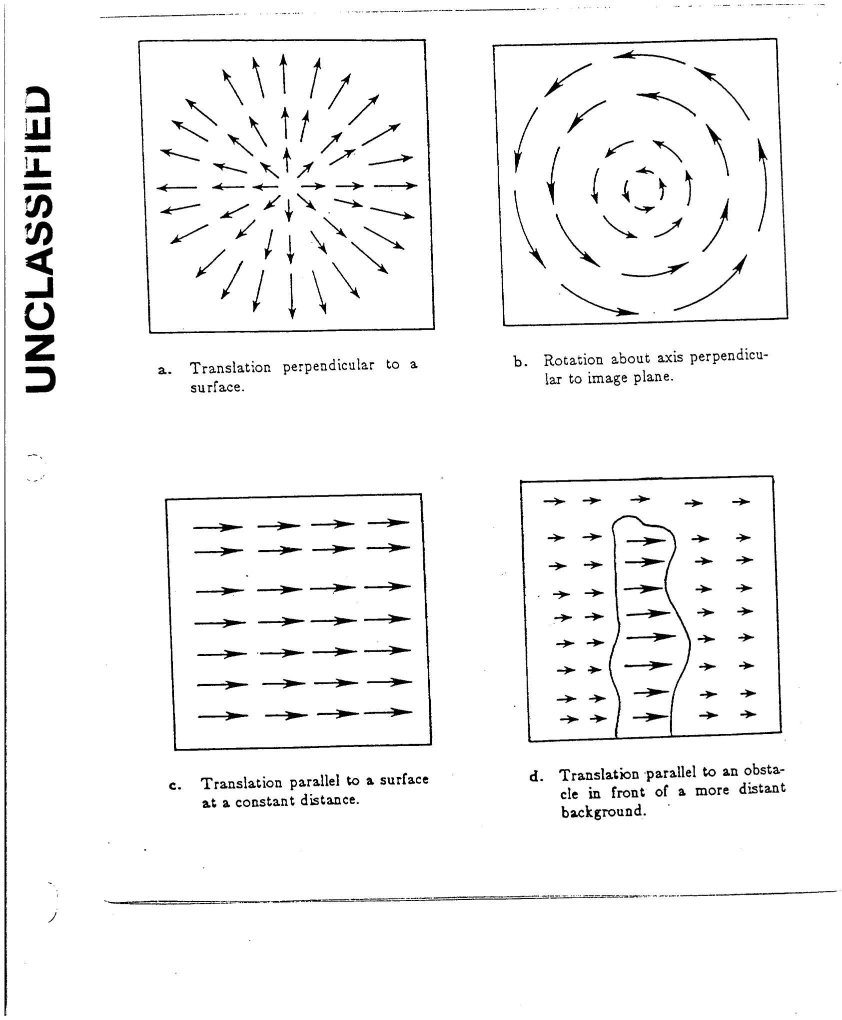


Figure 3.

A moving sensor induces an optic flow from stationary objects that depends on the objects' 3-D locations with respect to the direction of travel (see Figure 3 for examples). The ability to understand action in three dimensions based upon non-stereo motion, size, perspective, or occlusion cues is evident in ordinary cinematography. Depth is commonly dramatized by filming with the camera in motion.

Subconscious processes monitor this induced motion for its use in localization of non target objects required in reflex obstacle avoidance and path planning.

Most advanced vertebrates have the ability to maintain a visual fix on a target, whether the target is moving or not. The fixation is maintained through saccadic eye movements and smooth pursuit eye movements. When the fixating animal is also moving, additional information about the geometry of the environment is gained by the induced optic flow. This information is approximated in Figure 4.

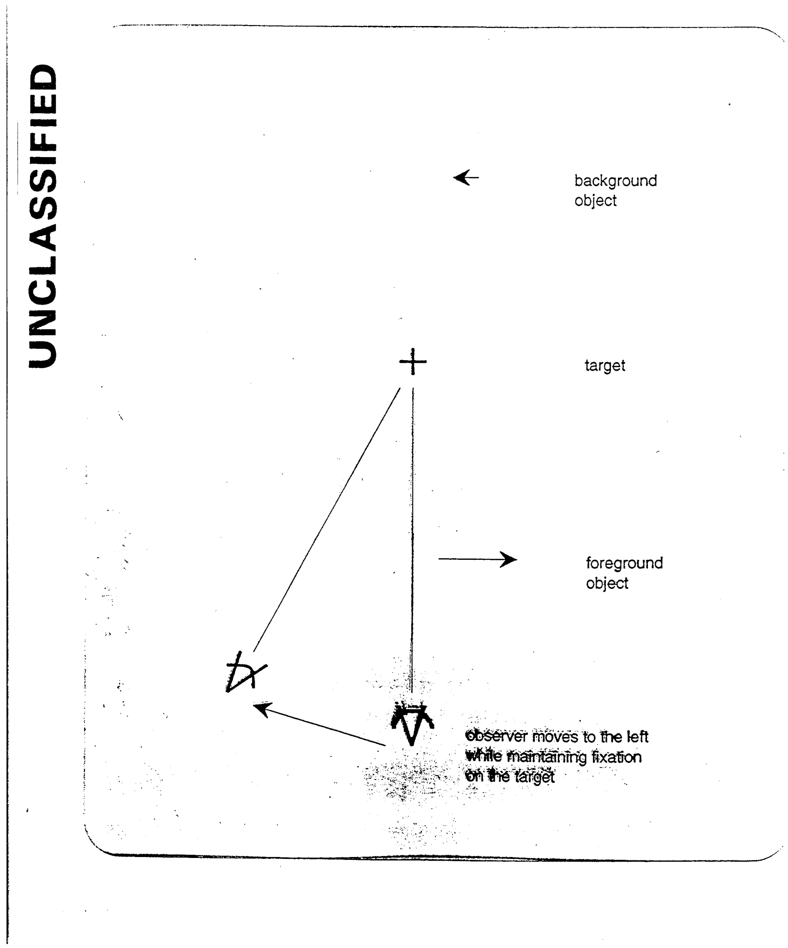


Figure 4.

Independent motion marks new targets

Animals use target motion as the principal cue for visual target acquisition. The superior colliculus, a midbrain nucleus responsible for selecting new visual targets, receives input from the motion detectors of the retina as well as from the cerebral cortex. Motion is a nearly irresistible factor in reflex control of visual attention. We are compelled to look at a target that moves uniquely, and while we may choose to look away, our attention is drawn back to it if it continues to exhibit erratic motion. Looking at a target means moving our eyes, head and body through saccades and smooth pursuit movements in the direction of the target so that the image of the target falls on the center of our retina (orienting reflex).

Motion segmentation mechanisms force attention to sources of unique motion (generally due to animate targets) and suppress conscious awareness of the consistent background motion (generally due to movements of the sensor).

Visual motion segmentation mechanisms permit target acquisition, tracking, and trailing.

Figure 5 shows a visually sensing robot acquiring, tracking, and trailing a walking human in a complex visual environment, using only visual motion segmentation for input.

Figure 5.

4) Motion reveals target behavior

When the target is in motion, the analysis of target motion is fundamental to the assessment of its behavior.

This is obvious. What it implies however is that we need mechanisms first to analyze or extract features from the motion flow, and second to integrate those features into patterns of motion (trajectories) that can evoke an appropriate response.

Intention is exposed in action.

5) The geometry of the vision system facilitates processing.

Animals generally have fixed sensor geometries, such as the distribution of receptors in the retina and the projection of their output onto the visual cortex.

In advanced vertebrates that use eye movements to scan for detailed information, the sensor geometry is modified to concentrate processing on the target region. This is the fovea of the retina. Peripheral input is compressed and used primarily for detection of new targets, based again on motion.

The primate visual system undergoes an approximate log-polar transformation from the photo receptors to the visual cortex. This transformation accomplishes data compression, committing a large part of the cortex to the processing of the central visual field (about 10 degrees visual angle), and a small portion to processing of the peripheral visual field (about 150 degrees on the horizontal). In addition, the transformation facilitates certain analyses of motion that are generally more relevant to an active vision system. For example, auto motion in the direction of the optical axis results in parallel flows on the computational plane.

Figure 6 shows the visual receptive fields, and the log-polar projection of the visual sensor employed by the robot in Figure 5.

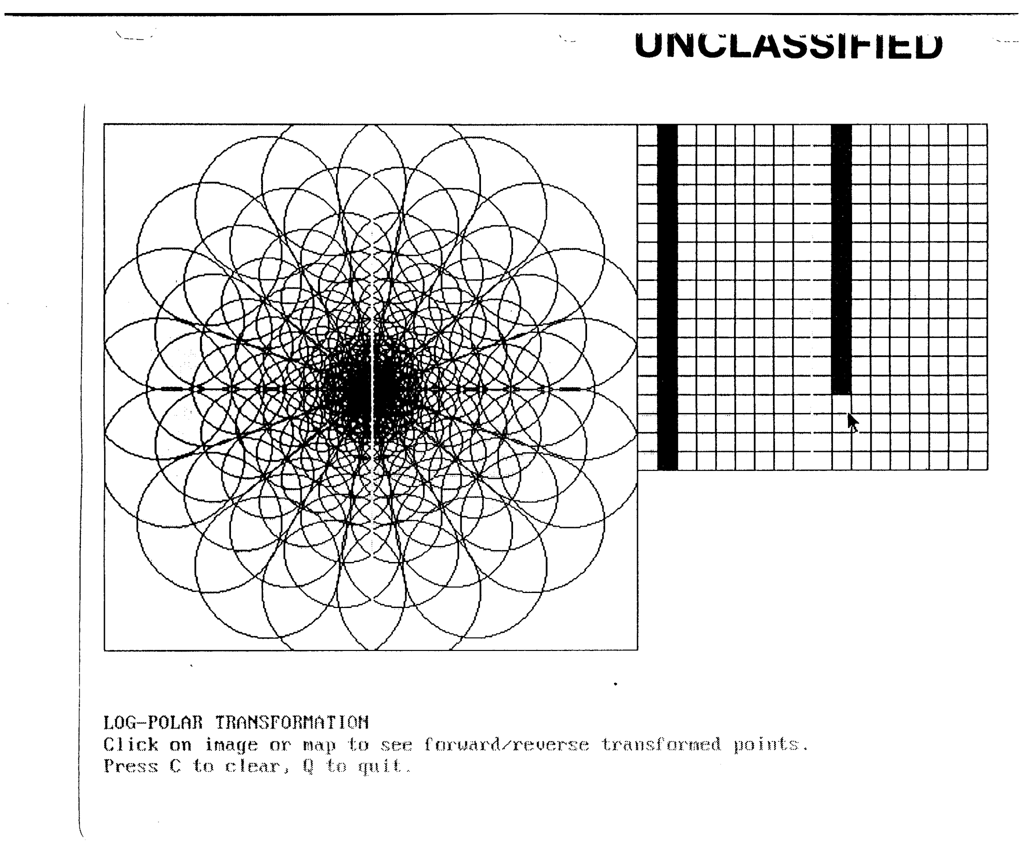


Figure 6.

The processing of sensor and motor information is closely related geometrically in the brain. The activation of a sensor field is likely to be associated with the activation of a motor field that controls muscles that further stimulates sensors projection to its

associated field. An example of this close correspondence is shown in Figure 7 in a sagital section of the human brain.

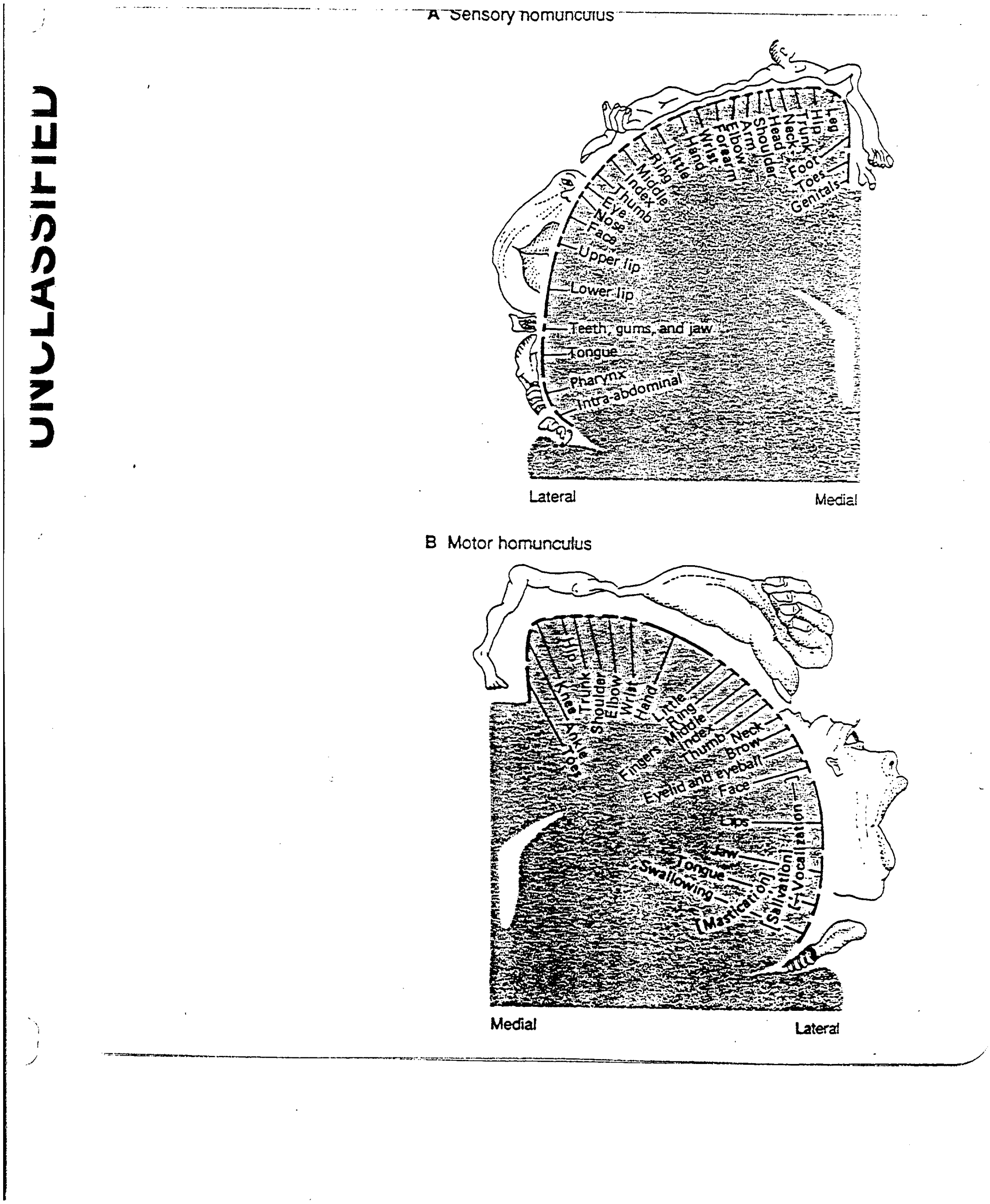


Figure 7.

6) Motion information can transform pattern information to achieve perceptual constancies.

Motion can also be used to transform extracted features to maintain alignment of predictions with subsequent observations, greatly reducing computational workload in

object recognition. While motion and pattern are known to be processed in parallel streams through the cortex, the two streams interact at several levels. Figure 8, from DeYoe and VanEssen (1988), summarizes the evidence.

The nervous system generally ignores constant input, whether or pattern or motion. Elementary pattern features, such as oriented lines, are most provocative when moving orthogonal to their preferred orientation.

1. Stationary features are ignored.

2. Oriented lines evoke stronger responses when moving
orthogonal to their preferred orientation.

3. Secondary and tertiary cortex contain higher percentages of cells that are direction specific.

4. Location specificity decreases while direction specificity increases with distance from primary sensory cortex.

UNCLASSIFIED

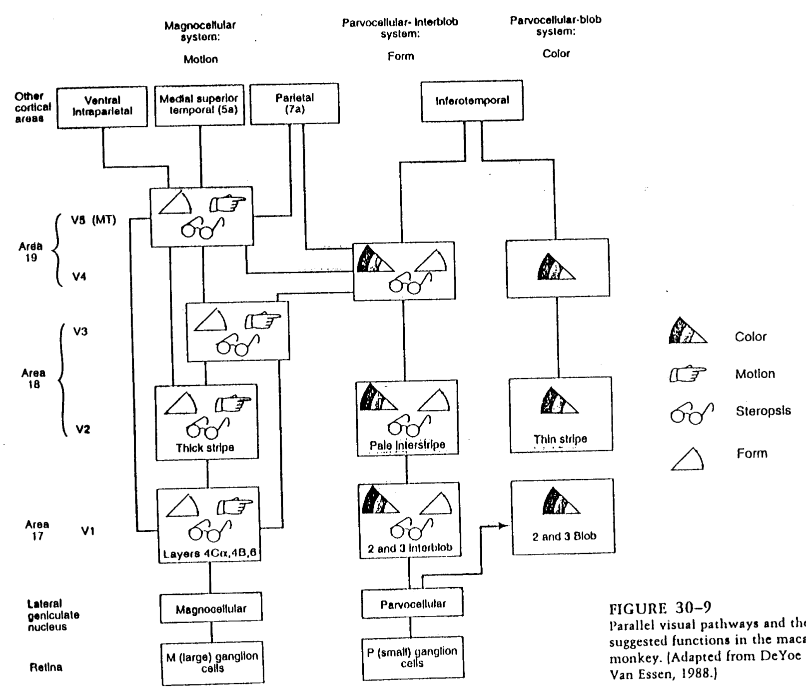


FIGURE 30–9
Parallel visual pathways and their suggested functions in the macaque monkey. (Adapted from DeYoe and Van Essen, 1988.)

Figure 8.

In the absence of visual input, the process can free run as the transformed features create new motion that leads to new transformations. What happens during a dream? The images move and often undergo unusual transformations. During a dream, eye movements occur (REM sleep) but are poorly organized. The reconstruction of images is a dynamic process, both creating motion and depending upon motion.

1. Dream images move on their own.

2. Dream images transform rationally and then decompose.

3. Vivid dreams are associated with poorly organized eye movements (REM sleep).

4. The reconstruction of images in a dream may involve the motion transformation of pattern features and the perception of new motion as a consequence. The process could then free-run.

5. Visual input during the waking state can justify motion-pattern interactions (reality testing).

7) Visual perception is an active process.

Also obvious.

The purpose of the central nervous system is not to dream, but to act. This perspective has been available in the neurobiological community at least since the time of Tolman (1932) and is frequently reiterated (Arbib, 1972; Pribram and Carlton, 1984; Roitblat, 1988, 1991; Varela, 1979). Its reciprocal, that the purpose of action is to perceive, is also voiced (Powers, 1973; Bandopadhay et al.,1986; Whitehead and Ballard,1990; Burt, 1988).

Experience allows discrimination.

Active perception is the application of control strategies to data acquisition based on the current state of data interpretation and the goal or task of the process (Bajcsy, 1988). Active perception occurs during the processes of autonomous sensor-effector control. Active perception is the execution of some behavior that results in the increased probability of encountering a specific stimulus. At a higher level, active perception attempts to satisfy a need for information. It can accomplish this by changing the relative perspective of the organism to its environment. Active perception is a means first to diversify contact with the environment and second to reduce distraction, improve the signal to noise ratio and reduce the computational requirements. Aloimonos et al. (1987) point out that problems that are ill-posed and nonlinear for a passive observer are well posed and linear for an active observer.

Uncertainty in the environment is the reason why active perception is required. An uncertain observer is evidenced by random behavior. Non-random behavior in a noisy environment is evidence for the success of active perception.

1. The real world contains uncertainty.

2. An uncertain agent acts randomly.

3. Non-random behavior is evidence for active perception.

4. Active perception is the application of experience to data collection.

5. Active perception increases the probability of finding a target.

6. Active perception reduces noise and computational requirements.

8) Reflex saccadic eye movements sample the environment.

Once a global search has acquired a target, a more detailed search performed by scanning mechanisms permits a logical sampling of target attributes, whether or not the target is itself moving. Target attributes compete for attention as do multiple targets observed from a distance. Figure 9 shows such a scan path produced by a human observer. The darkest blotches are the saccade target locations where the observer's eyes rested for approximately 0.5 sec prior to moving ballistically on to the next location.

Experience is gained through observing the order in the environment (correlations) produced during reflex reorientations to salient features of an object.

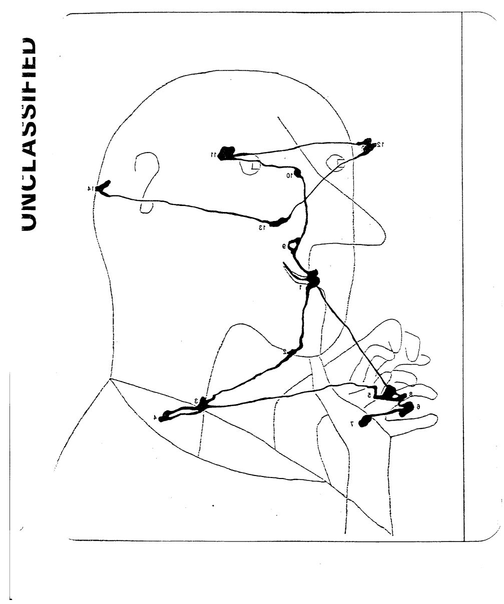


Figure 9.

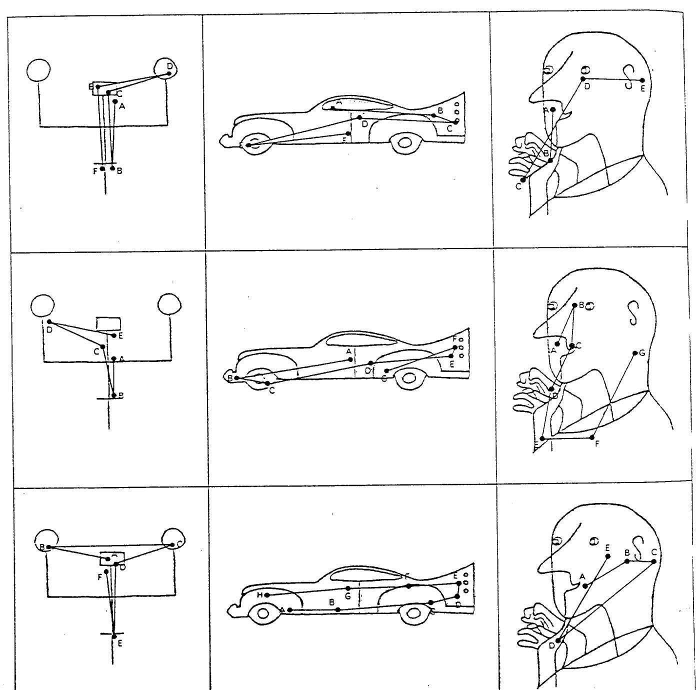
Smooth pursuit eye movement temporarily maintain the target on the high resolution fovea but are frequently interrupted by small saccades that continue to actively sample the geography of target attributes. The reader may easily verify this for himself by observing a moving automobile at 100 yards. The eyes will smoothly track the automobile, but will also jump from location to location on the body of the automobile to identify salient features.

9) Expectations drive search patterns over familiar targets.

After a period of observation when data collection is controlled primarily by reflex saccades, the vision system begins to anticipate the next saccade and preempts the reflex. Learned scan paths are the active processes of perception.

Rizzo et al. (1987) studied the fixation patterns of two patients with impaired facial recognition and learning and found an increase in the randomness of the scan patterns compared to controls, indicating that the cortex was failing to direct the search for relevant information with a degree of control that exceeded the attractive potential of the stimulus features.

UNCLASSIFIED



VARIETY IN SCAN PATHS is shown for three subjects and three pictures. Each horizontal row depicts the scan paths used by one subject for the three pictures. Vertically one sees how the scan paths of the three subjects for any one picture also varied widely.

Figure 10

Yarbus (1967) demonstrated the sensitivity of patterns of eye movements to the cognitive requirements of a visual search task. The regions of an image that were most often visited as a saccade target contained information relevant to the task. Without explicit task requirements, individuals had idiosyncratic scanpaths (Figure 10) suggesting that the sequence of saccades were determined not solely by the stimulus features, but by an interaction of stimulus features and an agenda brought to the task by the individual, that is, the individual demonstrated some expectations about the image to be viewed. Yarbus expressed this finding as "...people who think differently...see differently" (Yarbus, 1967, p. 211).

1. Experience allows anticipation of features that can interact with the target features and drive the scan path.

2. Learned scan paths are an active process of perception.

3. Brain damaged patients with poor face recognition have random scan paths.

4. Cognitive requirements (expectations) can influence a scan path.

10) Recognition is the verification of a prediction.

The verification of a prediction is the amplification of the current input that matches the reafferent activity, this process is similar to template matching or adaptive resonance theory of Carpenter and Grossberg (1987). An output results from an amplified input pattern as associated motor fields are recruited.

Recognition is a phase transition that changes the dynamic state of the system. It is not a point process or even a limit cycle, which are both maladaptive and incompatible with survival. The phase transition places the system in a new behavioral context, from which responses are deemed correct or incorrect by other observers.

In a study of scan paths and perception of the young woman/old woman ambiguous figure, Gale and Findlay (1983) found that fixation patterns correlated with the perception of the figure. The perception of an old woman (Figure 11) was accompanied by saccades that collected data on the mouth and nose of the figure (a vertical sequence of data acquisition) while the perception of a young woman was accompanied by saccades that collected data on the eye lash and ear (a horizontal data acquisition that missed the critical clues of the old woman in the figure).

1. Recognition builds from the accumulation of data that match expectations.

2. All high level brain states are normally transient.

3. Recognition may undergo a phase transition (hysteresis may be involved) after encountering data that mismatch the current bias.

2. The phase transition places the system in a different state with a different bias and different expectations.

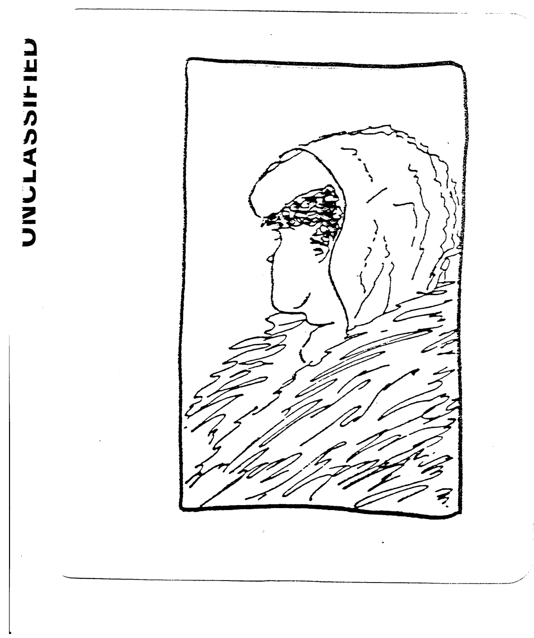


Figure 11.

11) Acquisition and use of information are inseparable processes.

In natural vision systems, the acquisition and use of information are not separable processes. Normally, irrelevant objects are ignored or quickly forgotten. Quite abstract two dimensional designs can gain significance for even some invertebrates if the design is correlated with the satisfaction of some vital need of the animal.

12) Animals learn environmental correlations to satisfy internal needs.

Animals learn not to please us, but to satisfy some internally sensed deficiency, such as hunger, thirst, restraint, sex, etc. The deficiency triggers an increase in neural activity (arousal) which is reduced in the course of satisfaction. This is diagramed in Figure 12.

1. Classical or Pavlovian conditioning is the model.

2. An internal deficiency is sensed such as hunger, thirst, cold, or pain;

3. Arousal is increased followed by activity;

4. The object of satisfaction is found followed by decreased arousal and activity.

5. Environmental features present during the change in arousal are associated with the sensed event that changed the arousal.

6. Thus we experience and anticipate rewards and punishments.

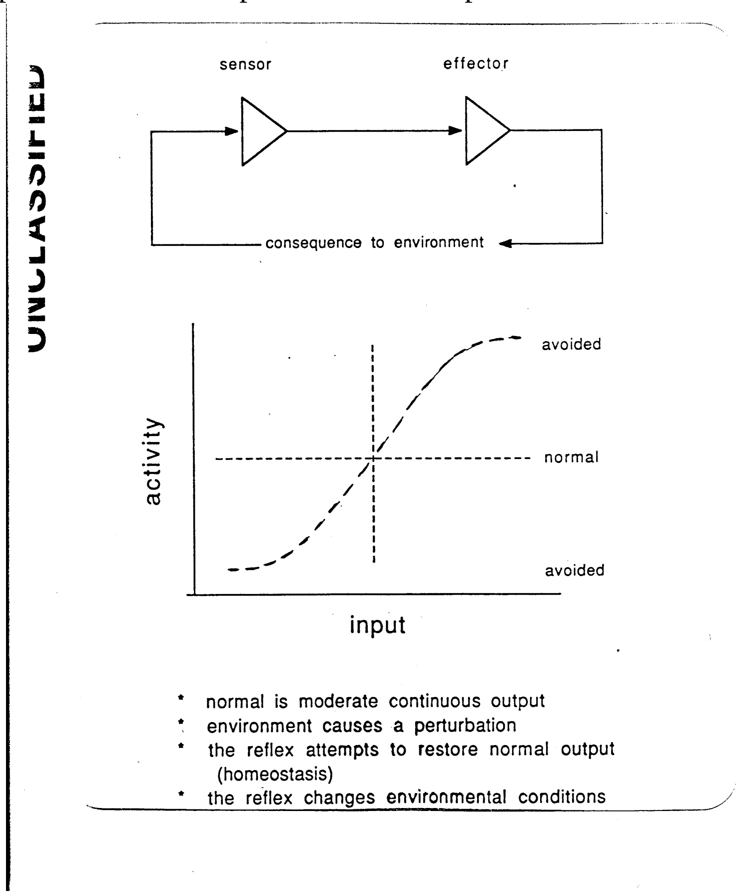


Figure 12

13) Machines can learn similarly if needs are appropriately defined and tested.

While the expected major benefit of using a machine vision system is freedom from the requirement to satisfy vital needs, the mechanisms involved in the acquisition and use of new information by a natural vision system are relevant to the development of analogous processes in an artificial vision system.

Motivation is generally ignored in machine learning. The learning process is controlled by an operator who determines when behavior is required, what behavior is required and which events are relevant for recall. In this scenario, the machine is not learning, instead the program parameters are being adjusted "on line". To approximate natural learning, a criterion for behavior must be sensed by the machine. Energy resources have been used (). When energy reserves drop, the activity of the machine is increased, when energy reserves are restored, the activity is reduced. Learning is accomplished in this protocol by correlating the motor output and sensory input present during the changes in activity and energy reserves. Events that lead to increases in activity (due to low energy reserves) are to be avoided, while events that lead to decreases in activity (due to restored energy reserves) are to be approached. This mechanism must allow for hysteresis, for activity itself will decrease reserves.

In every adaptive system, natural or synthetic, there are one or more reasons to change its structure and its input-output transfer function. In a supervised system, these reasons are exogenous. In an autonomous system, the reasons are endogenous. In a supervised autonomous system, the exogenous reasons are apparent to the supervisor, but they are effective only if they manipulate endogenous factors.

The appropriate selection of adaptation criteria in large part determines the success of adaptation. The mediation of the adaptation criteria is a biphasic process. Active network connections are strengthened when the output of the system contributes to the restoration of the criterion set-point values, and are weakened when it differs from those required values.

The experience of an artificial vision system with the types of information with which it must function, mediated by exogenous or endogenous reasons to change, allows the system to self-organize and determine, on its own, the relevant features, both in space and in time, that can be used to discriminate and respond appropriately to dynamic visual input.

1. Endogenous Motivation: energy reserves (useful in fielded systems), activity levels (optimize data collection per computational speed).

2. Exogenous Motivation: apply by manipulating one of the endogenous reflexes.

3. Strengthen associations between sensor fields when homeostatic set-points are approached. Weaken associations upon withdrawal from set-points.

4. The analogue of the arousal parameter may be the sensitivity of the perceptual system to phase transition.

5. Frequent changes in state with high arousal discourage discrimination learning.

14) Machine learning, following biological precedent, requires a reflex base, sensor preprocessing for feature definitions, abstract association matrices between sensor domains.

All behavior is built upon simple reflexes. One such reflex in shown in Figure 13. All complex behavior is achieved through the modulation of basic reflexes as shown in Figure 14.

Motivation is the result of a reflex increase in activity due to an interoceptor signalling some deficiency. The reflex base for behavior has several advantages: it provides self preserving behavioral defaults, it scales learning to the physical limits of the system, it keeps learning relevant, it connects elementary features with elementary motor responses.

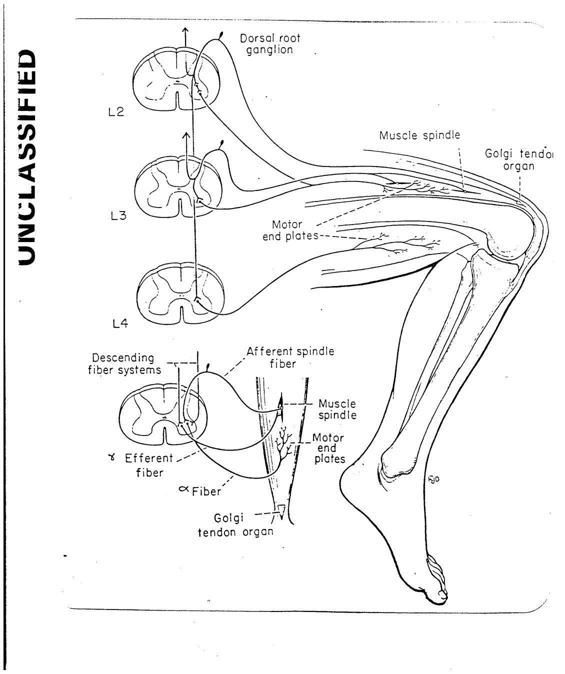


Figure 13.

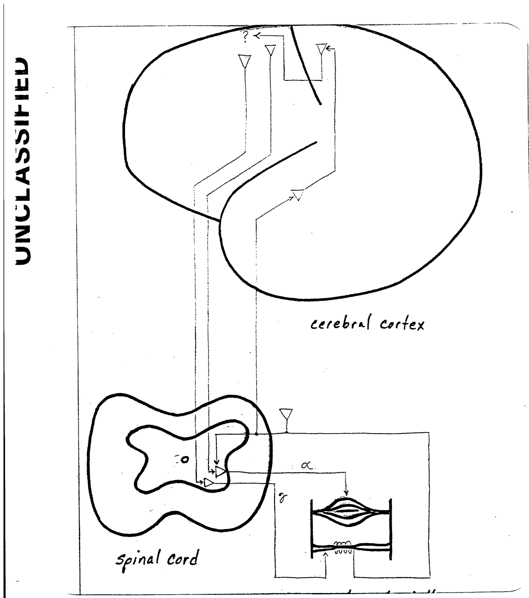


Figure 14.

Sensor preprocessing is a means to analyze input. Elementary features are made available for coding events. Multi-layer neural networks can learn discriminable coding, but at great computation cost. The natural neural system applies plasticity judiciously and not universally. No evidence of long tern plasticity in the spinal cord. Most functions of the brain stem, including the hypothalamus are species specific and innate. The organization of feature analyzers in primary cortex can be impaired with impoverished environments, but normal exposure yields similar results between individuals of a species. It is in the multi-sensory association cortex that neural responses cannot be predicted within a species. "Grandmother" cells apparently do not exist, rather the perception of one's grandmother is a spatial-temporal pattern of activity in larger numbers of cooperating neurons, resulting in the sequencing of multiple muscle groups. No single location in the nervous system contains a specific idea, or makes unilaterally a single decision. The natural neural network is a cooperative venture. Figure 15 shows an example of population coding.

Adaptation is correlated with visual capability in nature, and where we want to improve capability in our artificial systems, we should explore the mechanisms of adaptation and incorporate these into our artificial systems.

The appropriate motor output of an adaptive polymodal sensor association field follows from the dynamic reconstructions of the elementary sensory fields that accompanied the correct or successful behaviors.

1. Reflex base.

2. Multiple sensor systems with feature extraction and recomposition hierarchy.

3. Association matrices between high level features of different sensor modalities.

4. Topographical mapping of sensor features and motor mechanisms - for scan paths, voice production, teletype, etc.

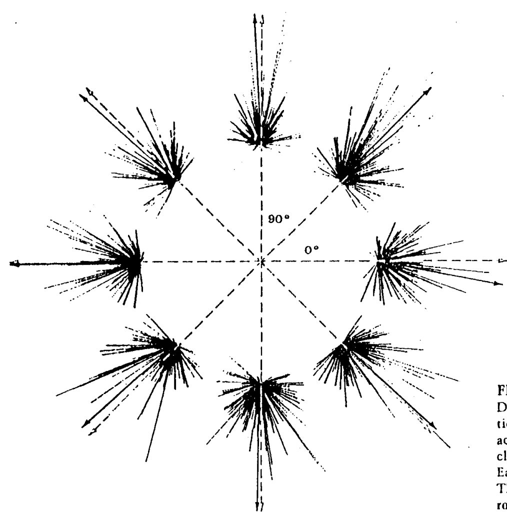


FIGURE 40–9
During movements in any given direction various cortical neurons with different preferred directions are active but the direction of the population vector closely matches that of the direction of movement. Each cluster represents the activity of one population. The directions of the population vectors (dashed arrows) closely match the direction of the targets.

Figure 15.

15) Neural networks, whether biological or artificial, self organize and select idiosyncratically relevant features for discrimination and prediction of environmental contingencies.

As designers of an artificial visual system, we can specify the decomposition of an image but this does not guarantee that the resulting features will be present in the target and obvious to the machine vision system. We could find that it takes less work to allow the machine itself to determine what is relevant. It could do so by simply selecting the features that make it through its filters at the time the critical decisions are required.

In the process of self-organizing to regularities in the environment, desired responses to classes of environmental conditions become probable. Such an increase in the probability in the scan path of a machine vision system with learning is shown in Figure 16.

1. It is difficult for the designer to anticipate what is relevant for a learning system.

2. Natural and artificial learning systems discover relevant features and correlations from the order in the environment as filtered by the systems experience based predispositions.

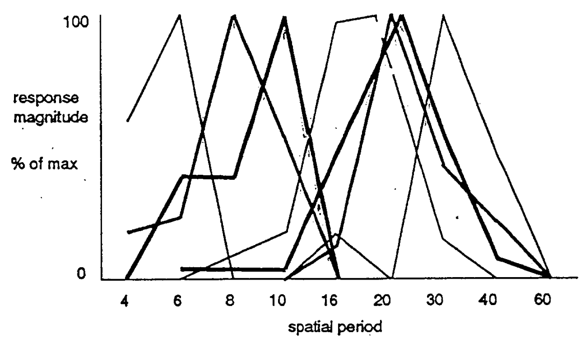


Figure 2. Activity profiles of complex elements taken from locations in the model cortex that receive projections with increasing eccentricity (a-g). The stimulus patterns were individual square wave gratings that were drifted slowly across the visual field. Note that the spatial period scale is nonlinear. The complex elements are sensitive to line orientation, spatial period and direction of motion relative to the surface of the cortex model.

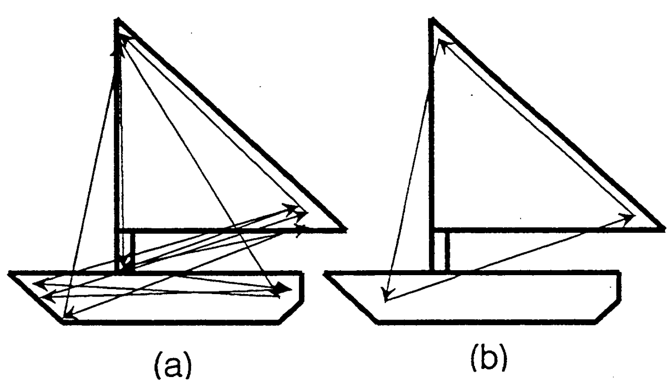


Figure 3. Scan paths of a naive network (a) and of a experienced network (b) to a static line drawing of boat. Arrows indicate direction of saccades. New target regions of the image were detected by small oscillations of the receptor surface, and selected competitively in the superior colliculus model network. Learning, based on experience scanning the image, provided a bias from the cortex to the superior colliculus that favored the regions of visual space from which the most likely (expected) saccade targets could be selected.

Figure 16.

16) Recommendations and Summary of Progress in machine vision at NRaD

The approach we advocate follows biological precedent and incorporates in its functional design low level deterministic specific responses to unspecific stimulus conditions (reflexes), monitored by accessory channels containing specific organizations for input pre-processing and output post-processing coupled by a large loosely differentiated matrix of adaptive processing elements, analogous to neurons or interneurons. The adaptation rules should be based on criteria relevant to the survival of the machine. The gross architecture of the artificial visual processing stages that we have implemented is shown in Figure 17. This architecture was used to learn the scan paths of Figure 16. Long-term adaptations (learning) were permitted only in the association cortex layer.

A large literature on both natural and artificial learning systems support this architecture and adaptation mechanism.

1. Emulate nature.

2. Include neurobiologists in design teams along with computer scientists.

3. Avoid historical biases.

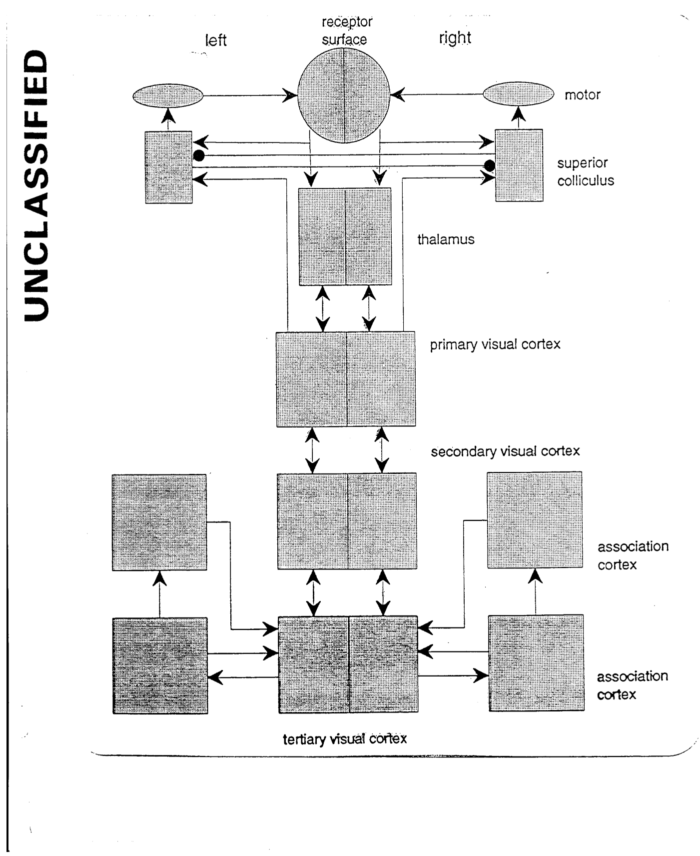


Figure 17.

We have available to date algorithms that emulate natural visual information processing. These algorithms perform 1) visual sensor to processing layer mapping that accomplishes data compression using a log-polar transformation (Blackburn, 1993a), 2) visual motion analysis of local activity in the log-polar domain (Blackburn and Nguyen, 1994b), 3) target acquisition and localization based on segmented motion (Blackburn and Nguyen, 1995), 4) feature analysis and re-synthesis by a hierarchical organization incorporating motion mediated transformations (Blackburn, 1993b), 5) adaptive associations of invariant spatio-temporal features and search behaviors (Blackburn, 1992), 6) cross modal adaptive sensor mapping as in Figure 18 (Blackburn and Nguyen, 1994b). The degree of maturity of these processes is inversely proportional to their order in the list.

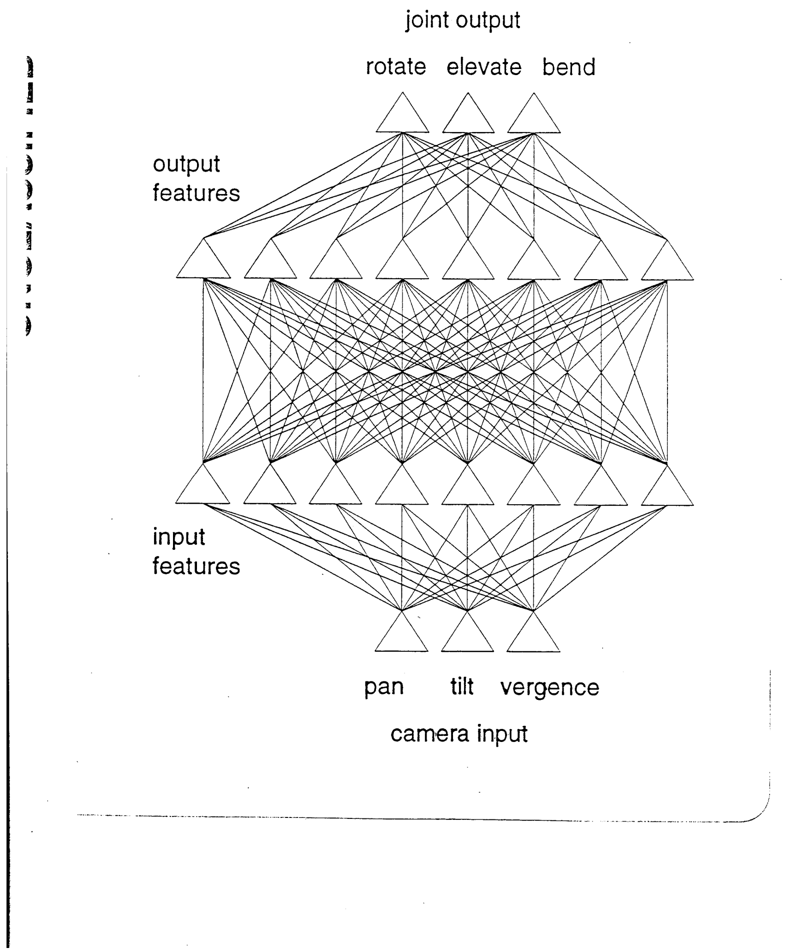


Figure 18.

References

Aloimonos, J., Weiss, I. and Bandyopadhyay, A. (1987) Active vision. CAR-TR-317, University of Maryland, Center for Automation Research

Bajcsy, R. (1988) Active perception. Proceedings of the IEEE, 76, 996-1005.

Bandopadhay, A., Chandra, C. and Ballard, D.H. (1986) Egomotion using active

vision. Proceedings Computer Vision and Pattern Recognition. I IEEE, 498-503.

Blackburn, M.R. (1992) Response invariance in an artificial visual-motor system. In IR-IED '92 Annual Report, NRaD TD 2412, Naval Command, Control and Ocean Surveillance Center RDT&E Division, San Diego, California, 9-26.

Blackburn, M.R., (1993a) A simple computational model of center-surround receptive fields in the retina. NRaD TD 2454, Naval Command, Control and Ocean Surveillance Center, RDT&E Division, San Diego, California.

Blackburn, M.R., (1993b) An algorithm for simple and complex feature detection: From retina to primary visual cortex. NRaD TD 2456, Naval Command, Control and Ocean Surveillance Center RDT&E Division, San Diego, California.

Blackburn, M.R. and Nguyen, H.G. (1994a) Learning in robot directed reaching: A comparison of methods. Proceedings of the 1994 Image Understanding Workshop, Monterey, CA. Nov 13-16, 781-788.

Blackburn, M.R. and Nguyen, H.G. (1994b) Autonomous visual control of a mobile robot. Proceedings of the 1994 Image Understanding Workshop, Monterey, CA. Nov 13-16, 1143-1150.

Blackburn, M.R. and Nguyen, H.G. (1995) Vision based autonomous robot navigation: Motion segmentation. Paper to be presented at the 28th ISATA Conference on Robotics, Motion, and Machine Vision in the Automotive Industries, Stuttgart, Germany, 18-22 September 1995.

Burt, P.J. (1988) Smart sensing within a pyramid vision machine. Proceedings of the IEEE, 76, 1006-1015.

Carpenter, G.A. and Grossberg, S. (1987) Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world. First International Conference on Neural Networks, San Diego, June 1987, II-737-745.

Gale, A.G. and Findlay, J.M. (1983) Eye movement pattern in viewing ambiguous figures. In Groner, R., Menz, C., Fischer, D.F. and Monty, R.A. (eds.) Eye Movements and Psychological Functions. Lawrence Erlbaum Assoc., Hillsdale, N.J., 145-168.

Powers, W.T. (1973) Behavior: The Control of Perception. Chicago, Aldine.

Pribram, K.H. and Carlton, E.H. (1984) Object perception. Manuscript.

Roitblat, H.L. (1988) Cognitive action theory: An action-based paradigm for connectionism. Neural Networks, 1, Supplement 1, 214.

Roiblat, H.L. (1991) Cognitive action theory as a control architecture. in J.-A. Meyer and S. Wilson (eds.) From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior. MIT Press, Cambridge, MA, 444-450.

Rizzo, M., Hurtig, R. and Damasio, A.R. (1987) The role of scan paths in facial recognition and learning. Annals of Neurology, 22, 41-45.

Tolman, E.C. (1932) Purposive behavior in Animals and Men. New York: Appleton-Century-Crofts.

Vallbo, A.B. (1994) Single-afferent neurons nad somatic sensation in humans. In M.S. Gassaniga (ed.) The Cognitive Neurosciences. The MIT Press, Cambridge, MA, 237-252.

van Essen, D.C. and Maunsell, J.H.R. (1983) Hierarchical organization and functional streams in the visual cortex. Trends in Neurosciences, 6, 1-6.

Varela, F.J. (1979) Principles of Biological Autonomy. New York: North Holland.

Whitehead, S.D. and Ballard, D.H. (1990) Learning to perceive and act. Technical Report 331. University of Rochester, Department of Computer Science.

Excellent single source for summaries of data, theories and models:

Gazzaniga, M.S. (ed.) The Cognitive Neurosciences. MIT Press, Cambridge, MA, 1995, 1446 pages.

Some of the viewgraphs used in this presentation were borrowed from figures printed in:

Kandel, E.R., Schwartz, J.H. and Jessell, T.M. (eds.) Principles of Neural Science, 3rd ed., Elsevier, New York, 1991, 1135 pages.